

# Propagation of chaos in neural fields

Jonathan Touboul

11, Place Marcelin Berthelot  
75005 Paris, France

e-mail: [jonathan.touboul@inria.fr](mailto:jonathan.touboul@inria.fr)

url: <http://www-sop.inria.fr/members/Jonathan.Touboul>

**Abstract:** We consider the problem of the limit of bio-inspired spatially extended neuronal networks including an infinite number of neuronal types (space locations), with space-dependent propagation delays modeling neural fields. The propagation of chaos property is proved in this setting under mild assumptions on the neuronal dynamics, valid for most models used in neuroscience, in a mesoscopic limit, the *neural-field limit*, in which we can resolve quite fine structure of the neuron's activity in space and where averaging effects occur. The mean-field equations obtained are of a new type: they take the form of well-posed infinite-dimensional delayed integro-differential equations with a nonlocal mean-field term and a singular spatio-temporal Brownian motion. We also show how these intricate equations can be used in practice to uncover mathematically the precise mesoscopic dynamics of the neural field in a particular model where the mean-field equations exactly reduce to deterministic nonlinear delayed integro-differential equations. These results have several theoretical implications in neurosciences we review in the discussion.

**AMS 2000 subject classifications:** Primary 60F99, 60B10; secondary 34C15.

**Keywords and phrases:** Mean-Field Limits, Propagation of chaos, Delayed Stochastic differential equations, Infinite-dimensional stochastic processes, Neural Fields.

## Introduction

Brain's activity results of the complex interplay of different cells, in particular neurons, electrical cells that manifest highly complex nonlinear behaviors characterized by the intense presence of noise. Neurons form large population assemblies at the scale of which emerge reliable and adapted responses to stimuli. Such local neural populations, often termed cortical columns, have a diameter of about  $50\mu m$  to  $1mm$ , contain few thousands to one hundred thousand neurons and are in charge of specific functions [25]. The interaction of several columns at different spatial locations allow processing of the complex sensory or cortical information and support brain functions. Such groups of cortical columns organize on the surface of the cortex and form spatially extended structures called *neural fields*, the activity of which is precisely at the scale most usual imaging techniques (e.g. EEG/MEG, optical imaging) record relevant phenomena, and also correspond to anatomical information revealed experimentally. A paradigmatic example is given by the primary visual cortex of certain mammals. In such cortical areas, neurons organize into columns responding preferentially to specific orientations in visual stimuli, and display specific connection patterns [19, 3]. The communication between neurons is characterized by a delay due to the transport of information through axons and to the typical time the synaptic machinery needs to transmit it. These delays have a clear role in shaping the neuronal activity, as established by different authors (see e.g. [7, 28]). In such structures, several highly populated columns interact, and the number of neurons in each column is orders of magnitude

higher than the number of columns (e.g., orientations) involved. A variety of important brain states rely on the coordinated behaviors of large neural assemblies and recently raised the interest of physiologists and computational neuroscientists. Among these, we shall cite the rapid complex answers to specific stimuli [32], decorrelated activity [11, 26], large scale oscillations [5], synchronization [20], and spatio-temporal pattern formation [14, 6].

The mathematical and computational analysis of the dynamics of neural fields relies almost exclusively on the use of heuristic models since the seminal work of Wilson, Cowan and Amari [1, 36]. These approach implicitly consider that averaging effects counterbalance the prominent noisy aspect of *in vivo* firing observed experimentally, and describe the mesoscopic cortical activity through a deterministic, scalar variable whose dynamics is given by integro-differential equations. This model was widely studied analytically and numerically, and successfully accounted for hallucination patterns, binocular rivalry and synchronization [22, 13]. Justifying these models starting from biologically realistic settings has since then been a great endeavor [4].

In this manuscript we undertake a rigorous analysis of neural fields. From the biological viewpoint, these are spatially extended cortical structures made of several highly populated neuronal ensembles (the neural *populations*) in charge of specific functions. From the mathematical viewpoint, neural fields are adequately described as the limit of a set of interacting nonlinear stochastic processes (generally governing the neuron's electrical potential and related variables) gathering into different homogeneous populations at specific locations on the cortex. Neurons in each population have similar dynamics and communicate with neurons of different populations depending on the respective positions of the populations on the cortex and after a specific time delay. In what we will call the *neural field limit*, both the number of neurons and the number of populations tend to infinity so that the populations completely cover a continuous space (a piece of cortex or a functional space).

This problem is evocative of statistical fluid mechanics and more generally interacting particle systems, and as such has been widely studied in mathematics and physics chiefly motivated by thermodynamics or fluid dynamics questions. In particular, the probability distribution of a typical set of particles in the limit where the total number of particles goes to infinity, and fluctuations around this limit where characterized for a number of models [24, 10, 31, 30, 29]. It was shown in several contexts that when considering that all particles have independent identically distributed initial conditions (*chaotic* initial conditions), then in the limit where the number of particles tends to infinity, the behavior of a few particles remains independent as time goes by, and all particles have the same probability distribution, which is the solution of a nonlinear Markov equation, often referred to as the *McKean-Vlasov* equation. The underlying biological problem motivates the introduction of a notion of spatial labeling of the (fixed) neurons, involving two mathematical aspects that were not covered in the literature. First is the fact that this induces the presence of infinitely many types of neurons (corresponding to the column neurons belong to), and second is the fact that since neurons communicate through the emission of electrical impulses transported at finite speed through the axons, space-dependent delays occur in the communication between two cells. These two aspects necessitate the development of the propagation of chaos theory towards infinite-dimensional functional settings that we aim at achieving in the present manuscript. We will show that in the neural field limit, the propagation of chaos property holds. Moreover, the activity is shown to converge in a certain sense towards the solution of a new object, a delayed integro-differential mean-field equation with space-dependent delays. This object has substantial differences with the usual McKean-Vlasov limits: beyond the presence of delays, the neural field limit regime is

at a mesoscopic scale where averaging effects locally to occur, but is fine enough to resolve brain's structure and its activity, resulting in the presence of an integral term over space. The speed of convergence towards the mean-field equations is quantified and involves two terms, one governing the averaging effect in each population and the second corresponding to the continuum limit. In the neural field regime, the limit equations are very singular, in particular trajectories are not measurable with respect to the space. These limits are very hard to analyze at this level of generality, and even numerical simulations are intricate in finite-populations settings [2]. However, in the type of models usually considered in the study of neural fields, namely the firing-rate model, we show in a companion article [33] that the behavior can be rigorously and exactly reduced to a system of deterministic integro-differential equations that are compatible with the usual Wilson and Cowan system in the zero noise limit. Noise intervenes in these equations a nonlinear fashion, fundamentally shaping in the macroscopic dynamics.

The paper is organized as follows. We start in section 1 by describing the mathematical setting of the study, abstracting classical relevant neuronal models that are specified and reviewed in appendix A, and more general models are considered in appendix B. We then analyze the integro-differential delayed McKean-Vlasov equations that will constitute our limit neural field equation in section 2 and demonstrate in particular their well-posedness, before addressing in section 3 the propagation of chaos property and convergence of the network equations towards the solutions of the mean-field equation. In section 4 we illustrate how this approach can be used in practice to analyze the effect of the parameters on the dynamics of the system in a particular example, reviewing some results of [33] afresh on a new example where noise, delays and spatial structure interact to shape the mesoscopic response of the neural field. The results of the mathematical analysis are then confronted to different recent experimental observations on collective dynamics of neural fields in the brain, and a few open problems of interest are discussed in the conclusion section 5.

## 1. Mathematical Setting

In all the manuscript, we are working in a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  endowed with a filtration  $(\mathcal{F}_t)_t$  satisfying the usual conditions. We consider a spatially extended network composed of  $N$  neurons, each neuron belonging to one of  $P(N)$  populations characterized by their locations  $(r_1, \dots, r_{P(N)}) \in \Gamma^{P(N)}$  on the cortex (or the feature space)  $\Gamma$ , a finite-dimensional compact set<sup>1</sup>. The state of each neuron  $i$  in the network is described by a  $d$ -dimensional variable  $X^i \in E := \mathbb{R}^d$ , typically corresponding to the membrane potential of the neuron and possibly additional variables such as those related to ionic concentrations and gated channels described in appendix A, and satisfy the network equations:

$$dX_t^{i,N} = \left( f(r_\alpha, t, X_t^{i,N}) + \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \sum_{p(j)=\gamma} \frac{1}{N_\gamma} b(r_\alpha, r_\gamma, X_t^{i,N}, X_{t-\tau(r_\alpha, r_\gamma)}^{j,N}) \right) dt + \sigma(r) dW_t^i. \quad (1)$$

---

<sup>1</sup>When considering  $\Gamma$  as the cortex, it will be a compact subset of  $\mathbb{R}^q$ ,  $q = 2$  or  $3$ , and when considering that populations are defined by the neuron's function, the shape of  $\Gamma$  can take different forms depending on the geometry of the feature space, for instance in the case of the primary visual area, neurons code for a preferred orientation of a visual stimuli that can be represented in the torus  $\Gamma = \mathbb{S}^1$ .

where  $f(r, t, x) : \Gamma \times \mathbb{R} \times E \mapsto E$  governs the intrinsic dynamics of each cell,  $(W_t^i)$  is a sequence of  $m$ -dimensional Brownian motions modeling the external noise and  $\sigma(r) : \Gamma \mapsto \mathbb{R}^{d \times m}$  a bounded and measurable function of  $r \in \Gamma$  modeling the level of noise at each space location, and  $b(r, r', x, y) : \Gamma^2 \times E^2 \mapsto E$  the interaction function of a neuron located at  $r'$  with voltage  $y$  on a neuron at location  $r$  with voltage  $x$ . The function  $\tau(r, r') : \Gamma^2 \mapsto \mathbb{R}^+$  is the interaction delay between neurons located at  $r$  and those at  $r'$  which is assumed to be a regular function of its two variables. We assume that all delays are bounded by a finite quantity  $\tau$ . The quantity  $r_\alpha$  is called the *location* of the population and  $\alpha$  is the population label. For a neuron  $i$  in the network, the population function  $p : \mathbb{N} \mapsto \mathbb{N}$  associates to a neuron  $i$  the population  $\alpha$  it belongs to. The number of neurons in each population in a network of size  $N$  defines a sequence of population size  $(N_1(N), \dots, N_{P(N)}(N))$  (we hence have  $\sum_{\gamma=1}^{P(N)} N_\gamma(N) = N$ ) corresponding to the number of neurons in population  $\gamma$  when the network size is equal to  $N$ . The number of populations  $P(N)$  and the number of neurons in each of these populations is assumed to be deterministic<sup>2</sup>. The interaction term presents a scaling factor  $\frac{1}{P(N)N_\gamma}$  ensuring the boundedness of the input received by neurons from population  $\gamma$  to the other populations, a biological fact related to the brain function and to the finiteness of the resources available for the synaptic transmission.

The different locations  $r_\gamma$  of the populations are related to the organization of the neurons on the space  $\Gamma$ . These locations are distributed according to a specific probability measure  $\lambda$  on  $\Gamma$ <sup>3</sup>. The locations of the  $P(N)$  populations,  $(r_1, \dots, r_{P(N)}) \in \Gamma^{P(N)}$ , are assumed to be randomly and independently drawn in  $\Gamma$  according to the probability  $\lambda(dr)$  in a different probability space  $(\Omega', \mathcal{F}', \mathbb{P}')$ . We will denote by  $\mathcal{E}$  the expectation over the realizations of the space locations  $(r_\alpha)$ .

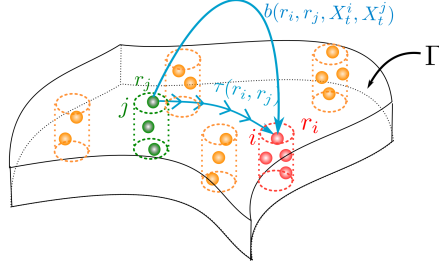


FIG 1. A typical architecture of neural fields: cylinders represent neural populations as cortical columns spanning across the cortex. Neuron  $i$  (red population at  $r_i \in \Gamma$ ) receives a spike from neuron  $j$  (green population at  $r_j \in \Gamma$ ) after a delay  $\tau(r_i, r_j)$  creating a current  $b(r_i, r_j, X_t^i, X_t^j)$ .

It is clear that the larger the number of populations, the smaller the mean number of neurons per populations. The number of populations will hence compete with the typical number of neurons per population and hence with averaging effects. In the present article, motivated by the fact that the number of neurons in each population is orders of magnitude larger than the number of populations (see e.g. [16]), we will make the following assumption, referred to as the

<sup>2</sup>It is easy to generalize to random population number and population size.

<sup>3</sup>In the example of the visual area V1,  $\lambda$  is the uniform measure on  $\mathbb{S}^1$

neural field limit:

$$\mathfrak{e}(N) := \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \frac{1}{N_\gamma(N)} \xrightarrow{N \rightarrow \infty} 0 \quad (2)$$

In the case of an infinite number of populations, this assumption ensures heuristically most populations are made of a diverging number of neurons<sup>4</sup>.

The parameters of the system are assumed to satisfy the following assumptions:

- (H1).  $f(r, t, \cdot)$  is uniformly  $K_f$  Lipschitz-continuous,
- (H2).  $b(r, r', \cdot, \cdot)$  is uniformly  $L$ -Lipschitz-continuous
- (H3). There exists a  $\tilde{K} > 0$  such that

$$|b(r, r', x, z)|^2 \leq \tilde{K}(1 + |x|^2)$$

- (H4). The drift satisfies uniformly in space ( $r$ ) and time ( $t$ ), the inequality:

$$|f(r, t, x)|^2 \leq C(1 + |x|^2)$$

- (H5). The drift, delay, diffusion and coupling functions are regular with respect to space variables  $(r, r') \in \Gamma^2$  (at least measurable, in practice generally assumed continuous).

Let us first state the following proposition ensuring well-posedness of the network system under the assumptions of the section:

**Proposition 1.** *Let  $(X_t^0)_{t \in [-\tau, 0]}$  a square integrable process with values in  $E^N$ . Under the assumptions of the section, there exists a unique strong solution to the network equations (1) with initial condition  $X^0$ , which is square integrable and defined for all times.*

The proof of this proposition is a direct application of Da Prato [8] as used by Mao [23], and essentially uses the same arguments as those of the proof theorem 2. The interested reader is invited to follow the steps of the demonstration of that theorem to prove proposition 1.

We are interested in the limit of such systems as the number of neurons  $N$  goes to infinity, under the neural field limit condition.

**REMARK.** *Let us briefly bring some results from the analysis of finite populations networks (i.e. the case where  $P(N)$  remains finite as  $N \rightarrow \infty$ ), which can be seen as a particular case of the current setting under the assumption that  $\lambda$  is a sum of Dirac masses. In that case, the neural field regime considers that the number of neurons in each population tends to infinity. Standard theory allows proving that the network converge towards  $P$  coupled McKean-Vlasov equations:*

$$d\bar{X}_t(r_\alpha) = f(r, t, \bar{X}_t(r_\alpha)) dt + \sigma(r_\alpha) dW_t^\alpha + \frac{1}{P} \sum_{\gamma=1}^P \mathbb{E}_{\bar{Z}}[b(r_\alpha, r_\gamma, \bar{X}_t(r_\alpha), \bar{Z}_{t-\tau(r_\alpha, r_\gamma)}(r_\gamma))] dt.$$

---

<sup>4</sup>If all populations have approximately the same number of neurons, each  $N_k(N)$  will be of the order  $N/P(N)$  and the condition (2) is satisfied when  $P(N) = o(N)$ . The condition also ensures the size of most populations tend to infinity. Indeed, for instance if all but one population contains just 1 neuron, the last population contains  $N - P(N)$  neurons, and the sum is equal to  $1 - 1/P(N) + 1/(P(N)(N - P(N))) \geq 1 - 1/P(N)$  which will not tend to zero.

where  $(W_t^\alpha)$  are  $P$  independent Brownian motions. This model can be seen as a discrete approximation of the continuous neural field. When the asymptotic number of populations is infinite, corresponding heuristically to refining the spatial discretization (or increasing the number of neurons) one is likely to face two main difficulties: (i) the network equations will involve an infinite number of independent Brownian motions, one for each space location, and (ii) it will involve a limit, as  $P$  goes to infinity, of a sum of the mean-field interaction terms (it is rather a simultaneous limit under the scaling property (2)). The infinite number of independent Brownian motions is not a technical artifact, but a fact related to the very nature of the problem: distinct neurons are driven by independent Brownian motions whatever their respective locations on the neural field  $\Gamma$ , and no continuity or measurability is to be expected.

In order to handle the first point, we introduce a particular object, the *spatially chaotic*<sup>5</sup> Brownian motion on  $\Gamma$ , a two-parameter process  $(t, r) \in \mathbb{R}^+ \times \Gamma \mapsto W_t(r)$  such that for any fixed  $r \in \Gamma$ , the process  $t \mapsto W_t(r)$  is a  $d$ -dimensional standard Brownian motion, and for  $r \neq r'$  in  $\Gamma$ , the processes  $W_t(r)$  and  $W_t(r')$  are independent. This process is relatively singular seen as a spatio-temporal process: in particular, it is not measurable with respect to the Borel algebra  $\mathcal{B}(\Gamma)$  of  $\Gamma$ . This object, defined as a collection of independent Brownian motions, clearly exists. More generally, in what follows, we will qualify a process  $\zeta_t(r)$  of spatially chaotic if the processes  $\zeta_t(r)$  and  $\zeta_t(r')$  are independent for any  $r \neq r'$ .

We will show that the network equations (1) satisfies the propagation of chaos property in the limit where  $N$  goes to infinity under the neural field assumption, and that the state of the network converges towards a very particular McKean-Vlasov equation involving a spatially chaotic Brownian motion. In detail, for almost all realizations of the spatial locations  $(r_\gamma, \gamma \in \mathbb{N})$  i.i.d. with law  $\lambda$ , the asymptotic law of neurons located at  $r$  in the support of  $\lambda$  will be measurable with respect to  $(\Gamma, \mathcal{B}(\Gamma))$  and converge towards the stochastic neural field mean-field equation with delays:

$$d\bar{X}_t(r) = f(r, t, \bar{X}_t(r)) dt + \sigma(r) dW_t(r) + \int_{\Gamma} \mathbb{E}_{\bar{Z}}[b(r, r', \bar{X}_t(r), \bar{Z}_{t-\tau(r, r')}(r'))] d\lambda(r') dt \quad (3)$$

where  $(W_t(r))_{t \geq 0, r \in \Gamma}$  is a spatially chaotic Brownian and the process  $(\bar{Z})$  is independent and has the same law as  $(\bar{X})$ . In other words, we will show that the law of the solution  $X_t(r)$ , noted  $m(t, r)(dy)$ , is measurable with respect to  $\mathcal{B}(\Gamma)$ , and that the mean-field equation can be expressed as the integro-differential McKean-Vlasov equation:

$$d\bar{X}_t(r) = f(r, t, \bar{X}_t(r)) dt + \int_{\Gamma} \int_E b(r, r', \bar{X}_t(r), y) m(t - \tau(r, r'), r')(dy) d\lambda(r') dt + \sigma(r) dW_t(r).$$

which will also be written, denoting  $\mathcal{E}_{r'}$  is the expectation with respect to the distribution of the population locations over  $\Gamma$  with distribution  $\lambda(\cdot)$ ,

$$d\bar{X}_t(r) = f(r, t, \bar{X}_t(r)) dt + \sigma(r) dW_t(r) + \mathcal{E}_{r'} [\mathbb{E}_{\bar{Z}}[b(r, r', \bar{X}_t(r), \bar{Z}_{t-\tau(r, r')}(r'))]] dt$$

Let us eventually give the Fokker-Planck equation on the possible density  $p(t, r, y)$  of  $m(r, t)$

---

<sup>5</sup>We use the term *chaotic* in the statistical physics sense as understood by Boltzmann's in his notion of molecular chaos ("Stoßzahlansatz").

with respect to Lebesgue's measure:

$$\begin{aligned} \partial_t p(t, r, x) = -\nabla_x \left\{ \left( f(r, t, x) + \int_{\Gamma} \int_E b(r, r', x, y) p(t - \tau(r, r'), r', y) d\lambda(r') \right) p(t, x) \right\} \\ + \frac{1}{2} \Delta_x [|\sigma(r)|^2 p(t, x)]. \end{aligned} \quad (4)$$

The mean-field equations (3) are of a new type: they resemble McKean-Vlasov equations but involve delays, spatially chaotic Brownian motions and an ‘integral over spatial locations’. This is hence a very unusual stochastic equation we need to thoroughly study in order to ensure that these make sense and are well-posed. The existence and uniqueness of solutions to these equations is addressed in section 2, and the proof of the propagation of chaos and convergence of the network equations towards the solutions of that equations is addressed in section 3.

REMARK. *Note that the setting considered here, though relatively general, can be further extended using locally Lipschitz-continuous drift and state-dependent diffusion functions as certain neuronal models require. Such refinements do not modify the results and the principles of the proofs, but induce an important increase of complexity in the presentation, and are commented in appendix B.*

## 2. Analysis of the mean-field equation

The mean-field equation (3) involves two unusual terms: a stochastic integral involving spatially chaotic Brownian motions and an integrated McKean-Vlasov mean-field term.

Let us start by discussing properties of stochastic integrals with respect to a spatially chaotic Brownian Brownian. Considering  $\Delta_t(r)$  a  $\mathcal{F}_t$ -progressively measurable process indexed by  $r \in \Gamma$  such that for any  $r \in \Gamma$  we have

$$\int_0^t \mathbb{E} [|\Delta_s(r)|^2] ds < \infty, \quad (5)$$

it is trivial to see that for any  $r \in \Gamma$ , the process  $N_t(r) := \int_0^t \Delta_s(r) dW_s(r)$  is a well defined, square integrable martingale with quadratic variation  $\int_0^t |\Delta_s(r)|^2 ds$ .

The possible solutions  $(\bar{X}_t(r))_{t,r}$  of the mean-field equation have a law belonging to the set of probability measures on the continuous functions of  $[-\tau, T]$  with values in the set of mappings of  $\Gamma$  in  $E$ . It is important to note at this point that similarly to the spatially chaotic Brownian motion, the solutions of the mean field equations are not measurable in  $(\Gamma, \mathcal{B}(\Gamma))$  since the solution considered at different space locations  $r$  and  $r'$  in  $\Gamma$ , namely  $X_t(r)$  and  $X_t(r')$ , are independent.

Though trajectories of spatially chaotic processes are non measurable, their probability distribution, defining a set of measures parametrized by  $r \in \Gamma$ , might be measurable. This is a necessary property to make sense of the mean-field equations. Handling this subtlety necessitates to thoroughly define the space in which we are working and where the mean-field equations are well-defined. We define  $\mathcal{Z}$  the set of random variables whose law is measurable with respect to  $\mathcal{B}(\Gamma)$  (the random variable itself is not assumed measurable with respect to  $\mathcal{B}(\Gamma)$ ). For a



random variable  $(Z(r))_{r \in \Gamma}$  in  $\mathcal{Z}$  with law  $p(r, dx)$  measurable, we define with a slight abuse of notations the  $\mathbb{L}_\lambda^2(\Gamma)$  norm on  $\Gamma$  by defining, for  $\hat{r}$  a  $(\Omega', \mathcal{F}', \mathbb{P}')$  random variable with law  $\lambda$ :

$$\|Z\|_{\mathbb{L}_\lambda^2(\Gamma)}^2 = \mathbb{E} [\mathcal{E}_{\hat{r}}[|Z(\hat{r})|^2]] = \int_{\Gamma} \int_E x^2 p(r, dx) d\lambda(r) \quad (6)$$

where  $\mathcal{E}$  denotes the expectation on  $\Omega'$ . This clearly defines a norm on random variables indexed by  $r \in \Gamma$ , when identifying processes that are  $\lambda \otimes \mathbb{P}$ -a.s. equal. We denote  $\mathbb{L}_\lambda^2(\Gamma)$  the set of random variables in  $\mathcal{Z}$  such that  $\|Z\|_{\mathbb{L}_\lambda^2(\Gamma)} < \infty$ .

EXAMPLE:

- (i). *The spatially chaotic Brownian motion at fixed time  $t$  has, for all  $r \in \Gamma$ , the law of a standard Brownian motion. This law, independent of  $r \in \Gamma$ , is hence measurable with respect to  $(\Gamma, \mathcal{B}(\Gamma))$ . Moreover, it belongs to  $\mathbb{L}_\lambda^2(\Gamma)$  and has a norm equal to  $t$ .*
- (ii). *Another example is given by the variable  $Z_T(r) = \int_0^T \Delta_s(r) dW_s(r)$  where  $\Delta$  is a function of  $\mathbb{R}^+ \times \Gamma$  measurable with respect to the  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{B}(\Gamma)$  and satisfying the condition  $\int_0^T \int_{\Gamma} |\Delta_s(r)|^2 d\lambda(r) ds < \infty$ . The thus defined variable is not measurable with respect to  $\mathcal{B}(\Gamma)$ , but belongs to  $\mathcal{Z}$  since this variable is a centered Gaussian process with measurable variance  $\int_0^T |\Delta_s(r)|^2 ds$ , hence the law of  $Z(r)$  is  $\mathcal{B}(\Gamma)$ -measurable. Eventually,  $Z \in \mathbb{L}_\lambda^2(\Gamma)$  with  $\|Z\|_{\mathbb{L}_\lambda^2(\Gamma)}^2 = \int_0^T \int_{\Gamma} |\Delta_s(r)|^2 d\lambda(r) ds$ .*

We extend this norm to processes with values in  $\mathbb{L}_\lambda^2(\Gamma)$ . For  $(Z_t(r))_{t \in [u, v]}$  a stochastic process indexed by  $r \in \Gamma$  such that the law of  $Z_t(r)$  is measurable with respect to  $\mathcal{B}(\Gamma)$ , we say that it belongs to  $\mathcal{M} := \mathcal{M}^2([u, v], \mathbb{L}_\lambda^2(\Gamma))$  if we have:

$$\|Z\|_{\mathcal{M}} := \mathcal{E}_{\hat{r}} \left( \mathbb{E} \left[ \sup_{s \in [u, v]} |Z_s(\hat{r})|^2 \right] \right) < \infty$$

and this quantity defines a norm on  $\mathcal{M}^2([u, v], \mathbb{L}_\lambda^2(\Gamma))$  where are identified the processes that are  $\lambda$ -a.s. and  $\mathbb{P}$ -a.s. equal for all times.

EXAMPLE:

- (i). *The spatially chaotic Brownian motion on  $[0, T]$  belongs to  $\mathcal{M}^2([0, T], \mathbb{L}_\lambda^2(\Gamma))$  and has a norm equal to  $T$  thanks to the classical property that the supremum of the Brownian motion has the law of the absolute value value of the Brownian motion.*
- (ii). *The process  $Z_t(r) = \int_0^t \Delta_s(r) dW_s(r)$  introduced above belongs to  $\mathcal{M} := \mathcal{M}^2([0, T], \mathbb{L}_\lambda^2(\Gamma))$  and, thanks to Burkholder-Davis-Gundy inequality, has a norm  $\|Z\|_{\mathcal{M}} \leq 4 \int_0^T \int_{\Gamma} |\Delta_s(r)|^2 d\lambda(r) ds$ .*

Now that these norms are introduced, we are in position to show the well-posedness of the mean-field equations:

**Theorem 2.** *For any  $(\zeta_t^0(r), t \in [-\tau, 0], r \in \Gamma) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  a square-integrable process, the mean-field equation (3) with initial condition  $\zeta^0$  has a unique strong solution on  $[-\tau, T]$  for any  $T > 0$ .*

*Proof.* As usually done for this kind of properties, we reduce the problem to the existence and uniqueness of a fixed point of a map  $\Phi$  acting on stochastic processes  $X$  in  $\mathcal{M}^2([-\tau, T], \mathbb{L}_\lambda^2(\Gamma))$



defined by:

$$\Phi(X)_t(r) = \begin{cases} \zeta_0^0(r) + \int_0^t f(r, s, X_s(r))ds + \int_0^t \sigma(r)dW_s(r) \\ \quad + \int_0^t \int_{\Gamma} \mathbb{E}_Z[b(r, r', X_s(r), Z_{s-\tau(r, r')}(r'))]d\lambda(r') ds, & t > 0 \\ \zeta_t^0(r), & t \in [-\tau, 0] \\ (Z_t) \stackrel{\mathcal{L}}{=} (X_t) \text{ independent of } (X_t) \text{ and } (W_t(\cdot)) \end{cases}$$

The first question we may ask is whether if this function is well defined, and if  $\Phi(X)$  defines a process that belongs to  $\mathcal{M}^2([-\tau, T], \mathbb{L}_{\lambda}^2(\Gamma))$ . The initial condition, Stieljes integral and stochastic integral with spatially chaotic Brownian motions are well defined as already seen. Moreover, under the assumption that  $X_t(r)$  has a law measurable with respect to  $\mathcal{B}(\Gamma)$ , each of these terms have a law measurable with respect to  $\mathcal{B}(\Gamma)$ . The mean-field term is slightly more complex. Let us denote by  $m(t, r, dy)$  the distribution of  $X_t(r)$ . This term can be written as:

$$\int_0^t \int_{\Gamma} \int_E b(r, r', X_s(r), y) m(s - \tau(r, r'), r', dy) d\lambda(r') ds$$

This quantity is well defined and has a distribution  $\mathcal{B}(\Gamma)$ -measurable with respect to  $r$  since we assumed  $m(t, r, dy)$  is measurable with respect to  $r$  as well as  $(r, r') \mapsto b(r, r', x, y)$  and  $(r, r') \mapsto \tau(r, r')$  (assumption (H5)). Let us now show square integrability of  $\Phi(X)$ . We have:

$$\begin{aligned} \|\Phi(X)\|_{\mathcal{M}_t^2}^2 &\leq 4 \left( \|\zeta_0^0\|_{\mathbb{L}_{\lambda}^2(\Gamma)}^2 + T C \int_0^T (1 + \|X_s\|_{\mathbb{L}_{\lambda}^2(\Gamma)}) + 4 \int_0^T \int_{\Gamma} |\sigma(r)|^2 d\lambda(r) ds \right. \\ &\quad \left. + T \tilde{K} \int_0^T (1 + \|X_s\|_{\mathbb{L}_{\lambda}^2(\Gamma)}^2) ds \right) \end{aligned} \quad (7)$$

which is finite since  $X \in \mathcal{M}_T^2$ .

We may hence iterate the map  $\Phi$ . We fix  $X$  a process in  $\mathcal{M}^2([-\tau, T], \mathbb{L}_{\lambda}^2(\Gamma))$  and build the sequence  $X^k$  by induction through the recursion relationship  $X^{k+1} = \Phi(X^k)$ . We aim at showing that these processes constitute a Cauchy sequence in  $\mathcal{M}^2([-\tau, T], \mathbb{L}_{\lambda}^2(\Gamma))$ , and we introduce the norm of the process up to time  $t$ :

$$\|X\|_{\mathcal{M}_t}^2 := \mathbb{E} \left[ \mathcal{E}_r \left( \sup_{s \in [-\tau, t]} |X_s(r)|^2 \right) \right] = \mathbb{E} \left[ \int_{\Gamma} \sup_{s \in [-\tau, t]} |X_s(r)|^2 d\lambda(r) \right]$$

We denote by  $(Z^k)$  a sequence of processes independent of the collection of processes  $(X^k)$  and having the same law. We study the norm  $\|X^{k+1} - X^k\|_{\mathcal{M}_T^2}$ . We decompose this difference into

the sum of three elementary terms as follows, for  $t \in [0, T]$  and  $r \in \Gamma$ :

$$\begin{aligned}
X_t^{k+1}(r) - X_t^k(r) &= \int_0^t \left\{ (f(r, s, X_s^k(r)) - f(r, s, X_s^{k-1}(r))) \right\} ds \\
&\quad + \int_0^t \int_\Gamma \left\{ \left( \mathbb{E}_Z[b(r, r', X_s^k(r), Z_{s-\tau(r, r')}(r'))] \right. \right. \\
&\quad \quad \left. \left. - \mathbb{E}_Z[b(r, r', X_s^{k-1}(r), Z_{s-\tau(r, r')}(r'))] \right) \right\} d\lambda(r') ds \\
&\quad + \int_0^t \int_\Gamma \left\{ \left( \mathbb{E}_Z[b(r, r', X_s^{k-1}(r), Z_{s-\tau(r, r')}(r'))] \right. \right. \\
&\quad \quad \left. \left. - \mathbb{E}_Z[b(r, r', X_s^{k-1}(r), Z_{s-\tau(r, r')}(r'))] \right) \right\} d\lambda(r') ds \\
&=: A_t(r) + B_t(r) + C_t(r)
\end{aligned}$$

We hence obviously have:

$$M_t^k := \|X^{k+1} - X^k\|_{\mathcal{M}_t^2}^2 \leq 3 \left( \|A\|_{\mathcal{M}_t^2}^2 + \|B\|_{\mathcal{M}_t^2}^2 + \|C\|_{\mathcal{M}_t^2}^2 \right)$$

We treat each term separately. We have:

$$\begin{aligned}
\|A\|_{\mathcal{M}_t^2}^2 &= \mathbb{E} \left[ \int_\Gamma \left( \sup_{s \in [0, t]} \left| \int_0^s f(r, u, X_u^k(r)) - f(r, u, X_u^{k-1}(r)) du \right|^2 \right) d\lambda(r) \right] \\
(\text{Cauchy-Schwartz}) &\leq TK_f^2 \mathbb{E} \left[ \int_\Gamma \left( \int_0^t |X_s^k(r) - X_s^{k-1}(r)|^2 ds \right) d\lambda(r) \right] \\
&\leq TK_f^2 \int_0^t \|X^k - X^{k-1}\|_{\mathcal{M}_s^2}^2 ds
\end{aligned}$$

which directly implies  $\|A\|_{\mathcal{M}_t^2}^2 \leq TK_f^2 \int_0^t M_s^{k-1} ds$ . The terms  $B_t$  and  $C_t$  can be controlled using the same techniques. Let us for instance treat the case of  $C_t$ . We have:

$$\begin{aligned}
\|C\|_{\mathcal{M}_t^2}^2 &= \mathbb{E} \left[ \int_\Gamma \sup_{s \in [0, t]} \left| \int_\Gamma \int_0^s \left( \mathbb{E}_Z[b(r, r', X_u^{k-1}(r), Z_{u-\tau(r, r')}(r'))] \right. \right. \right. \\
&\quad \left. \left. - \mathbb{E}_Z[b(r, r', X_u^{k-1}(r), Z_{u-\tau(r, r')}(r'))] \right) du d\lambda(r') \right|^2 d\lambda(r) \right] \\
(CS) &\leq t \int_{\Gamma^2} \int_0^t \mathbb{E} \left[ \mathbb{E}_Z \left[ \left| b(r, r', X_u^{k-1}(r), Z_{u-\tau(r, r')}(r')) - b(r, r', X_u^{k-1}(r), Z_{u-\tau(r, r')}(r')) \right|^2 \right] du \right] d\lambda(r) d\lambda(r') \\
(H2) &\leq t L^2 \int_{\Gamma^2} \int_0^t \mathbb{E} [|X_s^k(r) - X_s^{k-1}(r)|^2] ds d\lambda(r) d\lambda(r') \\
&\leq t L^2 \int_0^t \mathbb{E} [\|X_s^k(r) - X_s^{k-1}(r)\|_{\mathbb{L}_\lambda^2(\Gamma)}^2] ds = t L^2 \int_0^t M_s^{k-1} ds
\end{aligned}$$

The term  $B_t$  is treated exactly in the same manner and yields the inequality:

$$\|B\|_{\mathcal{M}_t^2}^2 \leq t L^2 \int_0^t M_s^{k-1} ds$$

All together we obtain, using the fact that for all  $k > 1$ ,  $t \in [-\tau, 0]$  and  $r \in \Gamma$  we have  $X_t^{k+1}(r) = X_t^k(r) = \zeta_t^0(r)$ :

$$M_t^k \leq K' \int_0^t M_s^{k-1} ds \quad (8)$$

with  $K' = 3T(K_f^2 + 2L^2)$ , readily implying:

$$M_t^k \leq (K')^k \int_0^t \int_0^{s_1} \dots \int_0^{s_{k-1}} M_{s_k}^0 ds_1 \dots ds_k \leq \frac{(K')^k t^k}{k!} M_T^0. \quad (9)$$

and  $M_t^0$  is finite since we assumed  $X \in \mathcal{M}_T^2$  and showed that  $\Phi(X) \in \mathcal{M}_T^2$ . Routine methods starting from inequality (9) using Benaymé-Tchebychev inequality and Borel-Cantelli lemma allow proving existence and uniqueness of fixed point for  $\Phi$  (see e.g. [27, pp. 376–377]), and that this fixed point is adapted and almost surely continuous. Moreover, this process has a measurable law with respect to  $(\Gamma, \mathcal{B}(\Gamma))$  as a limit of measurable laws, and since it satisfies  $\Phi(\bar{X}) = \bar{X}$ , it is a solution to equation (3). Let us eventually show that it belongs to  $\mathcal{M}_T^2$ . We stop the process at time

$$\tau_n = \inf\{t > 0, \|X\|_{\mathcal{M}_t^2} > n\}.$$

Using inequality (7) and the fact that  $\bar{X} = \Phi(\bar{X})$  we have:

$$\|\bar{X}\|_{\mathcal{M}_{t \wedge \tau_n}^2} \leq 4 \left( \|\zeta_0^0\|_{\mathbb{L}_\lambda^2(\Gamma)} + T(C + \tilde{K} + 4\|\sigma\|_{\mathbb{L}_\lambda^2(\Gamma)}) + T(C + \tilde{K}) \int_0^{t \wedge \tau_n} \|X\|_{\mathcal{M}_s^2} ds \right)$$

ensuring by Gronwall's lemma that  $\tau_n$  a.s. tends to infinity as  $n \rightarrow \infty$ , and eventually, letting  $n \rightarrow \infty$ , that the solution has a finite norm in  $\mathcal{M}_T^2$ .

Proving uniqueness of the solution using equation (8) is then folklore.  $\square$

Now that we proved strong existence and uniqueness of solutions for the mean-field equations, we now turn to showing that network equations indeed convergence in law towards this solution, and that the propagation of chaos occurs.

### 3. Limit in law and propagation of chaos

We are now in a position to prove the main result of the manuscript, namely the convergence in law of the solutions of the network equations (1) towards the equations (3), and the fact that the propagation of chaos property occurs. To this end, we consider that the network equations have chaotic initial conditions. In details, let  $(\zeta_t^0(r)) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  a spatially chaotic stochastic process, i.e. a stochastic process such that for any  $r \neq r'$ , the process  $(\zeta_t^0(r))$  is independent of  $(\zeta_t^0(r'))$ . We consider that the initial condition of different neurons in the network are independent, and the initial condition  $(\zeta_t^i) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  for neuron  $i$  in population  $\alpha$ , is equal to  $(\zeta_t^0(r_\alpha)) \in \mathcal{M}^2(\mathcal{C}_\tau)$ .

The classical coupling argument cannot be directly applied here. Indeed, the usual argument is based on the fact that we are able to define the solution of the mean-field equation through the use of the *same* Brownian motion and with the *same* initial condition as one of the neurons (or particles). This is no more the case because individual neurons are governed by finite-dimensional Brownian motions and the mean-field equation by a spatially chaotic Brownian

motion. Notwithstanding, an argument based on a slightly more subtle couplings holds. In details, let us consider neuron  $i \in \mathbb{N}$  of the network, in population  $\alpha$  at location  $r_\alpha \in \Gamma$ . Denote by  $(\tilde{W}_t^i)$  the Brownian motion governing the evolution of neuron  $i$  in the network and  $\zeta^i \in \mathcal{M}(\mathcal{C}_\tau)$  the initial condition of the network. We aim at defining a spatially chaotic Brownian motion  $W_t^i(r)$  on  $\mathbb{R}^{m \times d}$  such that the standard Brownian motion  $(W_t^i(r_\alpha))$  is equal to  $(\tilde{W}_t^i)$ , and proceed as follows. Let  $(W_t(r))_{t \in [0, T], r \in \Gamma}$  be a  $m \times d$ -dimensional spatially chaotic Brownian motions independent of the processes  $(\tilde{W}_t^j)$ . The processes:

$$\begin{cases} (W_t^i(r)) = (W_t(r)) & r \neq r_\alpha \\ (W_t^i(r_\alpha)) = (\tilde{W}_t^i) \end{cases}$$

is clearly a spatially chaotic Brownian motions, and will be used to construct a particular solution of the mean-field equations. In order to completely define a solution of the mean-field equations, we need to specify an initial condition, and aim at coupling it to the initial condition of neuron  $i$ . To this end, we define a spatially chaotic process  $(\tilde{\zeta}_t^0(r)) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  equal in law to  $(\zeta_t^0(r))$  and independent of  $\zeta_t^i$ , and define a coupled process  $(\zeta_t^{i,0}(r)) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  as:

$$\begin{cases} \zeta_t^{i,0}(r) = \tilde{\zeta}_t^0(r) & r \neq r_\alpha \\ \zeta_t^{i,0}(r_\alpha) = \zeta_t^i. \end{cases}$$

Here again, it is clear that this process is spatially chaotic, i.e. that for any  $r \neq r'$ , the processes  $\zeta_t^{i,0}(r)$  and  $\zeta_t^{i,0}(r')$  are independent, and that  $\zeta_t^{i,0}(r)$  has the law of  $\zeta_t^0(r)$ .

Now that these processes have been constructed, we are in a position to define the process  $(\bar{X}_t^i)$  as the unique solution of the mean-field equation (3), driven by the spatially chaotic Brownian motion  $(W_t^i(r))$  and with the spatially chaotic initial condition  $(\zeta_t^{i,0}(r))$ :

$$\begin{cases} d\bar{X}_t^i(r) = f(r, t, \bar{X}_t^i(r)) dt + \int_\Gamma \mathbb{E}_Z[b(r, r', \bar{X}_t^i(r), Z_{t-\tau(r, r')}(r'))] d\lambda(r') dt \\ \quad + \sigma(r) dW_t^i(r), & t \geq 0 \\ \bar{X}_t^i(r) = \zeta_t^{i,0}(r), & t \in [-\tau, 0] \\ (Z_t) \stackrel{\mathcal{L}}{=} (\bar{X}_t^i) \in \mathcal{M} \quad \text{independent of } (\bar{X}_t^i), (W_t^i(\cdot)) \text{ and } (B_t^i(\cdot, \cdot)) \end{cases}.$$

The same procedure applied for all  $j \in \mathbb{N}$  allows building a collection of independent stochastic processes  $(\bar{X}_t^j(r))_{j=1 \dots N} \in \mathcal{M}^2([-\tau, T], \mathbb{L}_\lambda^2(\Gamma))$  such that all neurons  $j$  in population  $\alpha$  have the same law as  $(\bar{X}_t(r_\alpha))$ . Let us denote by  $m(t, r)$  the probability distribution of  $\bar{X}_t(r)$  solution of the mean-field equation (3). As previously, the process  $(Z_t(r))$  generically denotes a process belonging to  $\mathcal{M}^2([-\tau, T], \mathbb{L}_\lambda^2(\Gamma))$  and distributed as  $m$ .

Let us fix  $l \in \mathbb{N}^*$  and  $(i_1, \dots, i_l)$  a collection of neuron indexes such respectively belonging to populations located at  $(r_1, \dots, r_k)$  (possibly identical). We now prove the almost sure convergence of a collection of processes  $(X_t^{i_k, N}, k = 1 \dots l)$  towards  $(\bar{X}_t^{i_k}(r_k), k = 1 \dots l)$ , implying its convergence of the law towards the chaotic distribution  $m(t, r_1) \otimes \dots \otimes m(t, r_k)$  as  $N$  goes to infinity. We start by proving this property for  $l = 1$  before extending that result to  $l > 1$ .

**Theorem 3.** *Let  $i \in \mathbb{N}$  a fixed neuron in population  $\alpha$ . Under the assumptions (H1)-(H5) and the neural field assumption (2), for almost all realizations of the population locations  $(r_\alpha, \alpha \in \mathbb{N})$ , the process  $(X_t^{i, N}, t \leq T)$  solution of the network equations (1) converges in law towards the*

process  $(\bar{X}_t(r_\alpha), t \leq T)$  solution of the mean-field equations (3) with initial condition  $(\zeta_t^0(r))$  and moreover, the speed of convergence is given by:

$$\mathcal{E} \left( \mathbb{E} \left[ \sup_{-\tau \leq s \leq T} |X_s^{i,N} - \bar{X}_s^i(r_\alpha)|^2 \right] \right) = O \left( \mathfrak{e}(N) + \frac{1}{P(N)} \right) \quad (10)$$

REMARK. We recall that  $\mathcal{E}$  denotes the expectation on the distribution of the space locations  $(r_k)_{k=1 \dots P(N)}$  and  $\mathfrak{e}(N) = \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \frac{1}{N_\gamma(N)}$ .

*Proof.* The proof is based on evaluating the distance  $\mathbb{E}[\sup_{-\tau \leq s \leq T} |X_s^{i,N} - \bar{X}_s^i|^2]$ , and breaking it into a few elementary, easily controllable terms. A substantial difference with usual mean-field proofs is that we need to prove a convergence in the infinite-dimensional space  $\mathbb{L}_\lambda^2(\Gamma)$ , and that the interaction term in networks equations consists of a sum over a finite number of populations, whereas the effective interaction term arising in the mean-field equation is an integral over  $\Gamma$ .

In all the demonstration, we will generically denote  $r_\beta \in \Gamma$  the location of population  $\beta \in \{1, \dots, P(N)\}$ . We use the following elementary decomposition (each line of the righthand side corresponds to one term of the decomposition,  $A_t(N) - E_t(N)$ ):

$$\begin{aligned} X_t^i - \bar{X}_t^i(r_\alpha) &= \int_0^t (f(r_\alpha, s, X_s^i) - f(r_\alpha, s, \bar{X}_s^i(r_\alpha))) ds \\ &+ \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} (b(r_\alpha, r_\gamma, X_s^i, X_{s-\tau(r_\alpha, r_\gamma)}^j) - b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), X_{s-\tau(r_\alpha, r_\gamma)}^j)) ds \\ &+ \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} (b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), X_{s-\tau(r_\alpha, r_\gamma)}^j) - b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), \bar{X}_{s-\tau(r_\alpha, r_\gamma)}^j(r_\gamma))) ds \\ &+ \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \left( \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), \bar{X}_{s-\tau(r_\alpha, r_\gamma)}^j(r_\gamma)) - \mathbb{E}_Z[b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), Z_{s-\tau(r_\alpha, r_\gamma)}(r_\gamma))] \right) ds \\ &+ \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \left( \mathbb{E}_Z[b(r_\alpha, r_\gamma, \bar{X}_s^i(r_\alpha), Z_{s-\tau(r_\alpha, r_\gamma)}(r_\gamma))] - \int_\Gamma \mathbb{E}_Z[b(r_\alpha, r', \bar{X}_s^i(r_\alpha), Z_{s-\tau(r_\alpha, r')}(r'))] d\lambda(r') \right) ds \\ &=: A_t(N) + B_t(N) + C_t(N) + D_t(N) + E_t(N) \end{aligned}$$

Due to the exchangeability of neurons belonging to the same population, the probability distribution of these terms do not depend on the particular neuron  $i$  considered but only on the population it belongs to. The terms  $A_t(N)$ ,  $B_t(N)$  and  $C_t(N)$  involve the Lipschitz continuity of the functions involved, the term  $D_t(N)$  correspond to averaging effects (mean-field limit) at single populations levels and the term  $E_t(N)$  corresponds to the continuous limit. The terms  $A_t(N)$  through  $C_t(N)$  are treated using the Lipschitz continuity of the functions involved. Using

Cauchy-Schwarz (CS) inequalities, we easily obtain:

$$\begin{aligned}\mathbb{E}[\sup_{0 \leq s \leq t} |A_s(N)|^2] &\leq K_f^2 T \int_0^t \mathbb{E}[\sup_{-\tau \leq u \leq s} |X_u^{i,N} - \bar{X}_u^i(r_\alpha)|^2] ds \\ \mathbb{E}[\sup_{0 \leq s \leq t} |B_s(N)|^2] &\leq T L^2 \int_0^t \mathbb{E}[\sup_{-\tau \leq u \leq s} |X_u^{i,N} - \bar{X}_u^i(r_\alpha)|^2] ds \\ \mathbb{E}[\sup_{0 \leq s \leq t} |C_s(N)|^2] &\leq T L^2 \int_0^t \max_{j=1 \dots N} \mathbb{E}[\sup_{-\tau \leq u \leq s} |X_u^{j,N} - \bar{X}_u^j(r_{p(j)})|^2] ds\end{aligned}$$

Let us for instance treat the case of  $B_t(N)$ :

$$\begin{aligned}\mathbb{E}[\sup_{0 \leq s \leq t} |B_s(N)|^2] &= \frac{1}{P(N)^2} \mathbb{E}\left[\sup_{0 \leq s \leq t} \left| \sum_{\gamma=1}^{P(N)} \int_0^s \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} \left( b(r_\alpha, r_\gamma, X_u^{i,N}, X_{u-\tau(r_\alpha, r_\gamma)}^{j,N}) \right. \right. \right. \\ &\quad \left. \left. \left. - b(r_\alpha, r_\gamma, \bar{X}_u^i, X_{u-\tau(r_\alpha, r_\gamma)}^{j,N}) \right) du \right|^2\right] \\ (\text{CS}) &\leq \frac{T}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} \mathbb{E}\left[ \left| b(r_\alpha, r_\gamma, X_s^{i,N}, X_{s-\tau(r_\alpha, r_\gamma)}^{j,N}) - b(r_\alpha, r_\gamma, \bar{X}_s^i, X_{s-\tau(r_\alpha, r_\gamma)}^{j,N}) \right|^2 \right] ds \\ ((\text{H2})) &\leq T L^2 \int_0^t \mathbb{E}\left[ |X_s^{i,N} - \bar{X}_s^i|^2 \right] ds \\ &\leq T L^2 \int_0^t \mathbb{E}\left[ \sup_{-\tau \leq u \leq s} |X_u^{i,N} - \bar{X}_u^i|^2 \right] ds.\end{aligned}$$

The mean-field term  $D_t(N)$  involve the difference between an empirical mean of a function of processes and an expectation term, and all have bounded second moment thanks to theorem 2 and assumption (H3). We have using (CS) inequality:

$$\mathbb{E}[\sup_{0 \leq s \leq t} |D_s(N)|^2] \leq \frac{T}{P(N)} \sum_{\gamma=1}^{P(N)} \int_0^t \mathbb{E}\left[ \left| \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} b(r_\alpha, r_\gamma, \bar{X}_s^i, \bar{X}_{s-\tau(r_\alpha, r_\gamma)}^j) - \mathbb{E}_Z[b(r_\alpha, r_\gamma, \bar{X}_s^i, Z_{s-\tau(r_\alpha, r_\gamma)}^\gamma)] \right|^2 \right] ds$$

and hence involves an expectation of type:

$$\begin{aligned}\mathbb{E}\left[ \left| \frac{1}{N_\gamma} \sum_{j=1}^{N_\gamma} \Theta(\bar{X}_s^i, \bar{X}_s^j) - \mathbb{E}_Z[\Theta(\bar{X}_s^i, Z_s^\gamma)] \right|^2 \right] &= \frac{1}{N_\gamma^2} \sum_{k,l=1}^{N_\gamma} \mathbb{E}\left[ (\Theta(\bar{X}_s^i, \bar{X}_s^j) - \mathbb{E}_Z[\Theta(\bar{X}_s^i, Z_s^\gamma)])^T \right. \\ &\quad \left. \cdot (\Theta(\bar{X}_s^i, \bar{X}_s^k) - \mathbb{E}_Z[\Theta(\bar{X}_s^i, Z_s^\gamma)]) \right]\end{aligned}$$

where  $\Theta(x, y) = b(r_\alpha, r_\gamma, x, y)$ . Routine methods allow to show that all the terms of the sum corresponding to indexes  $j$  and  $k$  such that the three conditions  $j \neq i$ ,  $k \neq i$  and  $j \neq k$  are satisfied are null. One simple way to show this property consists in writing the expectations as integrals with respect to the measure  $m(t, r_\alpha)$  and observing that all terms annihilate. Therefore, there are no more than  $3 N_\gamma$  non-null terms in the sum (in the case  $\alpha = \gamma$  there are just  $N_\gamma$

non-null terms), and moreover, all of these terms are uniformly bounded. The terms related to indexes  $j = k \neq i$  satisfy the inequality:

$$\begin{aligned} \mathbb{E} \left[ \left| \Theta(\bar{X}_s^i, \bar{X}_s^j) - \mathbb{E}_Z[\Theta(\bar{X}_s^i, Z_s^\gamma)] \right|^2 \right] &\leq 2 \mathbb{E} \left[ \left| \Theta(\bar{X}_s^i, \bar{X}_s^j) \right|^2 + \left| \mathbb{E}_Z[\Theta(\bar{X}_s^i, Z_s^\gamma)] \right|^2 \right] \\ &\leq 2 \left\{ \tilde{K}(1 + \mathbb{E}[|\bar{X}_s^i|^2]) + \mathbb{E} \left[ \left| \mathbb{E}_Z[\sqrt{\tilde{K}(1 + |\bar{X}_s^i|^2)}] \right|^2 \right] \right\} \\ &\leq 4\tilde{K}(1 + C'(s)) \end{aligned}$$

with  $C'(s)$  given by theorem 2. The terms related to the cases  $j = i$  (or symmetrically  $k = i$ ) are bounded by the same constant, since we have for all  $k$  such that  $p(k) = \alpha$ , using Cauchy-Schwartz inequality. We note  $C = 4\tilde{K}(1 + C'(T))$ . We hence conclude that:

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} |D_s(N)|^2 \right] \leq T^2 C \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \frac{3N_\gamma - 1}{N_\gamma^2} \leq 3T^2 C \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \frac{1}{N_\gamma} = 3T^2 C \mathbb{E}(N)$$

It hence only remains to control the term  $E_t(N)$  corresponding to the difference between an integral over the space  $\Gamma$  weighted by the density  $d\lambda(r)$  and a sum, weighted by  $1/P(N)$  of the same integrand at  $P(N)$  discrete values  $(r_\gamma) \in \Gamma^{\mathbb{N}}$  independently drawn in  $\Gamma$  with the probability density  $d\lambda(r)$ . This sum hence resembles to a Monte-Carlo approximation of the integral term, and we now show that our sums over populations converges for almost all choice of  $(r_\gamma) \in \Gamma^{\mathbb{N}}$  towards the integral, using an argument similar to the one we just used to control  $D_t(N)$ . In details, we show that  $\mathcal{E}(\mathbb{E}[\sup_{0 \leq s \leq t} |E_s(N)|^2])$  converges towards 0, using the same method as used for the convergence of the mean-field term. Let us denote for the sake of compactness of notations  $F(s, r, r')$  the expectation  $\mathbb{E}_Z[b(r, r', \bar{X}_s^i(r), Z_{s-\tau(r, r')}(r'))]$ .

We have:

$$\mathcal{E}(\mathbb{E}[\sup_{0 \leq s \leq t} |E_s(N)|^2]) \leq T \int_0^t \mathcal{E} \left( \mathbb{E} \left[ \left| \int_\Gamma \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} F(s, r_\alpha, r_\gamma) - \mathcal{E}_{r'}[F(s, r_\alpha, r')] \right|^2 ds \right] \right)$$

Similarly to what was done for the term  $D_t(N)$ , since  $\mathcal{E}_{r'}[F(s, r_\alpha, r')]$  is precisely the expectation of  $F(s, r_\alpha, r_\gamma)$  under the law of  $r_\gamma$  over which the sum is taken, developing the squared sum into a double sum over populations (say,  $\gamma$  and  $\gamma'$ ), it is easy to show that, because of the independence of the  $r_\gamma$ , that all terms that do not correspond to  $\gamma = \gamma'$ ,  $\gamma = \alpha$  or  $\gamma' = \alpha$  vanish, leaving less than  $3P(N)$  possibly non-null terms, and these terms are uniformly bounded. Indeed, for  $r_\gamma = r_{\gamma'}$  (the case  $r_\gamma = r_\alpha$  is treated in the same manner), we have:

$$\begin{aligned} \mathcal{E}(\mathbb{E}[|F(s, r_\alpha, r_\gamma) - \mathcal{E}_{r'}[F(s, r_\alpha, r'_\gamma)]|^2]) &\leq 2\mathcal{E}(\mathbb{E}[|F(s, r_\alpha, r_\gamma)|^2] + |\mathcal{E}_{r'}[F(s, r_\alpha, r'_\gamma)]|^2]) \\ &\leq 2\mathcal{E}(\mathbb{E}[|F(s, r_\alpha, r_\gamma)|^2] + \mathcal{E}_{r'}[|F(s, r_\alpha, r'_\gamma)|^2]) \\ &\leq 4\tilde{K}(1 + C(s)) \end{aligned}$$

implying eventually that:

$$\mathcal{E}(\mathbb{E}[\sup_{0 \leq s \leq t} |E_s(N)|^2]) \leq \frac{4T^2 \tilde{K}}{P(N)} (1 + C(T)).$$



All together, we hence have:

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq s \leq t \wedge \tau_U} |X_s^{i,N} - \bar{X}_s^i(r_\gamma)|^2 \right] &\leq K' \int_0^{t \wedge \tau_U} \max_{j=1 \dots N} \mathbb{E} \left[ \sup_{-\tau \leq u \leq s} |X_u^{j,N} - \bar{X}_u^j(r_{p(j)})|^2 \right] ds \\ &\quad + C_1 \mathfrak{e}(N) + \mathbb{E} \left[ \sup_{0 \leq s \leq t} |E_s(N)|^2 \right] \end{aligned}$$

valid for all  $i \in \mathbb{N}$ , hence we have:

$$\begin{aligned} \mathcal{E} \left[ \max_{i=1 \dots N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i(r_\gamma)|^2 \right] \right] \\ \leq K' \int_0^{t \wedge \tau_U} \mathcal{E} \left[ \max_{j=1 \dots N} \mathbb{E} \left[ \sup_{-\tau \leq u \leq s} |X_u^{j,N} - \bar{X}_u^j(r_{p(j)})|^2 \right] \right] ds + C_1 \mathfrak{e}(N) + \frac{C_2}{P(N)} \end{aligned}$$

where  $K' = 4T(K_f^2 + 2L^2)$ ,  $C_1 = 12T^2C$  and  $C_2 = 16T^2\tilde{K}(1 + C(T))$  neither depend upon  $N$  nor in the particular neuron considered. By Gronwall's inequality, we hence obtain:

$$\mathcal{E} \left[ \max_{j=1 \dots N} \mathbb{E} \left[ \sup_{-\tau \leq s \leq t \wedge \tau_U} |X_s^{j,N} - \bar{X}_s^j(r_\gamma)|^2 \right] \right] \leq \left( C_1 \mathfrak{e}(N) + \frac{C_2}{P(N)} \right) \frac{e^{K'T}}{K'}$$

which ends the proof.  $\square$

**Corollary 4.** *Let  $l \in \mathbb{N}^*$  and fix  $l$  neurons  $(i_1, \dots, i_l) \in \mathbb{N}^*$ . Under the assumptions of theorem 3, the law of  $(X_t^{i_1,N}, \dots, X_t^{i_l,N}, -\tau \leq t \leq T)$  converges towards  $m_t(r_{p(i_1)}) \otimes \dots \otimes m_t(r_{p(i_l)})$  for almost all realization of the population locations  $(r_\alpha, \alpha \in \mathbb{N})$ .*

*Proof.* We have:

$$\begin{aligned} \mathcal{E} \left( \mathbb{E} \left[ \sup_{-\tau \leq t \leq T} \left| (X_t^{i_1,N}, \dots, X_t^{i_l,N}) - (\bar{X}_t^{i_1}, \dots, \bar{X}_t^{i_l}) \right|^2 \right] \right) \\ \leq \sum_{k=1}^l \mathcal{E} \left( \mathbb{E} \left[ \sup_{-\tau \leq t \leq T} \left| X_t^{i_k,N} - \bar{X}_t^{i_k} \right|^2 \right] \right) \\ \leq l \left( C_1 \mathfrak{e}(N) + \frac{C_2}{P(N)} \right) \frac{e^{K'T}}{K'} \end{aligned}$$

which tends to zero as  $N$  goes to infinity, hence the law of  $(X_t^{i_1,N}, \dots, X_t^{i_l,N}, -\tau \leq t \leq T)$  converges towards that of  $(\bar{X}_t^{i_1}, \dots, \bar{X}_t^{i_l}, -\tau \leq t \leq T)$  whose law is equal by definition to  $m(t, r_{p(i_1)}) \otimes \dots \otimes m(t, r_{p(i_l)})$ .  $\square$

**IMPORTANT REMARK.** *The speed of convergence towards the mean-field equation is hence governed by  $\mathfrak{e}(N)$  and  $1/P(N)$ . In the case of a finite number of populations, the speed of convergence is hence driven by the size of the smallest population. In the infinite population case, the speed of convergence towards the mean-field limit is a balance between the averaged number of neurons in each population through the term  $\mathfrak{e}(N)$ , and the total number of populations through the term  $1/P(N)$ . The first term quantifies the speed at which averaging effects occur in the network, and is related to the averaged inverse number of neurons in each population. The other term controls the convergence of the interaction*

related to all populations towards an effective interaction term given by an integral over  $\Gamma$  of mean-field interactions, i.e. convergence of finite-populations networks towards their continuous limit. For networks with homogeneous population sizes,  $\mathfrak{e}(N)$  will be approximately equal to  $P(N)/N$ . The optimal network size ensuring the fastest convergence in that case hence corresponds to  $P(N) \sim \sqrt{N}$  (minimizing the functional  $x \mapsto P(x)/x + 1/P(x)$ ), and in that case the convergence will be in  $1/\sqrt{N}$ , and we conjecture that this speed of convergence is optimal (though we did not achieve to prove it). This convergence speed is hence very slow compared to finite-size networks and usual mean-field limits in which the speed of convergence is of order  $1/N$ .

#### 4. Neural Fields Equations in Action

It is folklore to notice that McKean-Vlasov limits have dynamics very complex to analyze. Very refined methods are generally set up to analyze the behavior of the system in the mean-field limit, such as entropy methods or spectral methods (see e.g. [34]). This statement could be even more true in our spatialized context, and the present, general approach might appear to be bound to remain formal.

Fortunately, for relevant neuroscience applications, it happens that solutions of these equations are not out of reach. This is the topic of a companion article [33] where is considered networks of firing-rate neurons (see appendix A.2), the neuronal model usually considered for neural fields analysis. Let us briefly review here the main results of that article and concretely use the proposed approach to analyze the dynamics of a simple network.

Considering firing-rate neurons, we show in [33] that the solutions of the mean-field equations are Gaussian processes when the initial condition also is (and equilibria are Gaussian), and that their mean  $M(r, t)$  and standard deviation  $v(r, t)$  (fully describing the process since the covariance is a simple function of these two quantities in that case) reduce to the set of deterministic delayed integro-differential equations:

$$\begin{cases} \partial_t M(r, t) &= -\frac{1}{\theta(r)} M(r, t) + I(r, t) \\ &+ \int_{\Gamma} J(r, r') F(r', M(r', t - \tau(r, r')), v(r', t - \tau(r, r'))) \lambda(r') dr' \\ \partial_t v(r, t) &= -\frac{2}{\theta(r)} v(r, t) + \sigma(r)^2 \end{cases} \quad (11)$$

where  $F(r, x, y)$  denote the expectation of  $S(r, U)$  for  $U$  a Gaussian random variable of mean  $x$  and variance  $y$ , and can be made explicit for particular choices of sigmoids  $S$ . These equations are consistent with the heuristically derived extremely widely used Wilson-Cowan models for finite-populations neural assemblies [35, 36] in the limit where noise levels vanish. These equations are shown to be well-posed, and address the existence of spatially homogeneous solution in law (i.e. solutions whose law do not depend on  $r \in \Gamma$ ). The analysis of the moment equations are shown to grant access to the dynamics of the network, and this is developed in details for some biologically motivated models. In that article, the choice of the parameters, driven by biological constraints, did not reveal any qualitative effect of the delays on the solutions except during transient phases.

In order to illustrate how the use of the present approach can be used to uncover the dynamics of the neural field, we proceed to the analysis of single population network with inhibitory interactions (i.e. negative interactions), a case that was not treated in [33] and which will turn out show a particularly rich variety of behaviors as a function of delays.

To this end, let us fix the parameters of the system. We consider  $\Gamma = \mathbb{S}^1$  the 1-dimensional torus, and  $\lambda$  the uniform distribution on it. We consider that  $S(r, x) = \int_0^{gx} e^{-x^2/2}/\sqrt{2\pi} =: \text{erf}(gx)$ ,  $\theta(r) = 1$  and  $\sigma$  independent of  $r$ , and can easily show by integration by parts that  $F(x, y) = \text{erf}(gx/\sqrt{1+g^2y})$ . We further fix  $J(r, r') = \bar{J}e^{-|r-r'|/\delta}$  ( $\delta$  represents the typical connectivity length in the neural field) and  $\tau(r, r') = |r - r'|/c + \tau_s$  ( $c$  represents the speed of transmission in the neural field and  $\tau_s$  the typical transmission time of the synapse).

Since  $F(0, y) = 0$  for any  $y \in \mathbb{R}$ , the Gaussian solutions with zero mean and standard deviation  $\sigma^2/2$  are stationary solutions of the system that are spatially homogeneous in law. Characterizing the stability of this solution consists in analyzing the characteristic roots equation of the linearized system around the spatially homogeneous stationary solution. Computing the eigenvalues of the integral convolution operator similarly to [33, Section 3.1], we obtain the *dispersion relationship*:

$$\xi + 1 = F'_0 \frac{e^{-\xi\tau_s}(1 - e^{-(\frac{1}{\delta} + \frac{\xi}{c})})}{\frac{1}{\delta} + \frac{\nu}{c} + i2\pi k}$$

for  $k \in \mathbb{Z}$  and  $F'_0 = \frac{g}{\sqrt{1+g^2\nu_0}} \frac{1}{\sqrt{2\pi}}$ . The spatially homogeneous equilibrium is stable if and only if all solutions  $\xi$  to the dispersion relationship (characteristic roots) have negative real parts. A Turing bifurcation point is defined by the fact that there exists an integer  $k$  such that  $\Re(\xi) = 0$ . It is said to be *static* if at this point  $\Im(\xi) = 0$ , and *dynamic* if  $\Im(\xi) = \omega_k \neq 0$ . In that latter case, the instability is called Turing-Hopf bifurcation, and generates a global pattern with wavenumber  $k$  moving coherently at speed  $\omega_k/k$  as a periodic wavetrain.

Possible Turing Hopf bifurcations hence arise when there exists  $\omega_k > 0$  such that:

$$i\omega_k + 1 = F'_0 e^{-i\omega_k\tau_s} Z_k(\omega)$$

with  $Z_k(\omega) = \frac{(1 - e^{-(\frac{1}{\delta} + \frac{i\omega}{c})})}{\frac{1}{\delta} + \frac{\nu}{c} + i2\pi k}$ , which yields bifurcation curves (parametrized by  $\omega$ ) in the parameter space:

$$\begin{cases} \sigma^2 &= \frac{2}{g^2} \left( -1 + \frac{\bar{J}^2 g^2 |Z_k(\omega)|^2}{2\pi(1 + \omega^2)} \right) \\ \tau_s &= \frac{-\arctan(\omega) + \text{Arg}(F'_0 Z_k(\omega)) + 2m\pi}{\omega} \end{cases}$$

This provides a curve of Turing Hopf bifurcations corresponding to transitions from stationary independent solutions to perfectly synchronized independent solutions, as displayed in figure Fig. 2 and 3. In Fig. 2 we display the bifurcation curve in the parameter space  $(\sigma, \tau_s)$  for a specific set of parameters. This curve has a convex shape. Small enough delays hence correspond to stationary solutions. Increasing delays yields periodic activity, which disappears as noise is increased. This example shows the importance of delays in the qualitative dynamics of the neural field. The typical connectivity length also shapes the qualitative dynamics of the neural field, as shown in Fig. 3. This variety of behaviors correspond to bifurcations corresponding to a wavenumber  $k = 0$ , and correspond to spatially homogeneous solutions. Non-trivial spatial structures can be searched for considering non-spatially homogeneous initial conditions. In this case, a number of complex spatio-temporal behaviors can appear, such as the metastable polychronization shown in Fig. 3 where the neural field splits into two clusters oscillating in antiphase during very long transient periods before a sudden synchronization of the whole neural field.

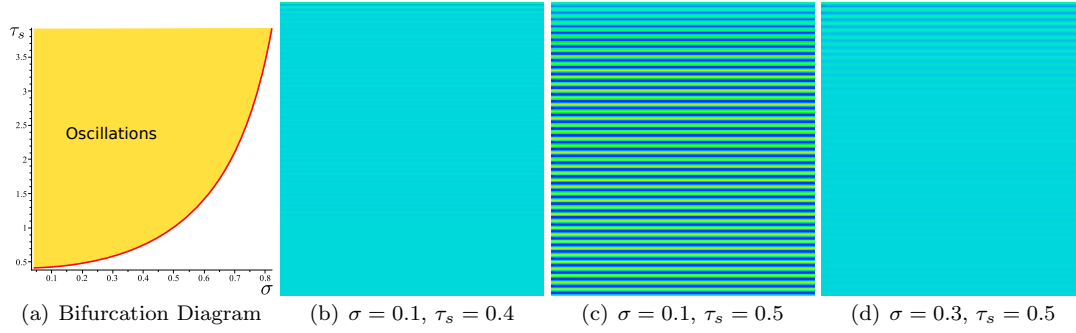


FIG 2. Turing-Hopf bifurcations and delay-induced synchronization,  $J = -3$ ,  $g = 3$ ,  $\delta = c = 1$ . (a): bifurcation diagram, shows a transition from stationary to periodic activity as delays are increased (b)→(c). When noise is increased, synchronization is lost (c)→(d). (b)-(d): spatio-temporal dynamics as a function of space (abscissa) and time (ordinate)

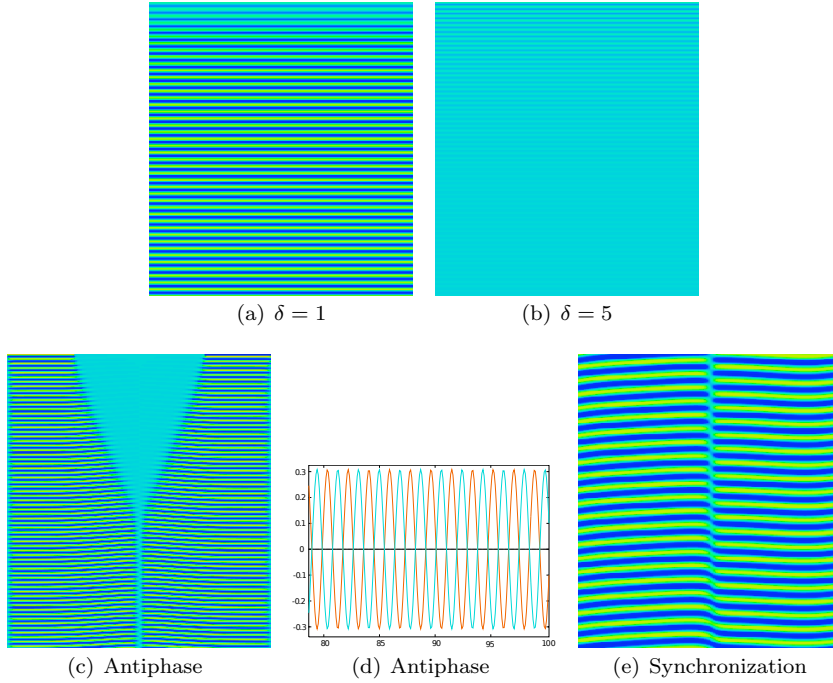


FIG 3. Spatial effects: increasing  $\delta$  destroys the synchronization. For  $\delta = 1$  (case (c) of Fig. 2), choosing non-spatially homogeneous initial conditions yields to complex situations, for instance an antiphase synchronization during long transients (bottom row). (c):  $t \in [0, 200]$ , (d): orange:  $M(t, 0.1)$ , blue:  $M(t, 0.9)$ , black:  $M(t, 0.5)$ , (e):  $t \in [600, 650]$ . The synchronization becomes visually perfect for times above 1500.

## 5. Discussion

In this paper, we addressed the problem of the asymptotic behavior of networks composed of a large number of neuronal assemblies in a particular asymptotic regime, the *neural-field limit*. We took into account a number of specificities relevant to neuronal dynamics: intrinsic noise at the level of each neuron, the spatial structure and propagation delays. We demonstrated that for a relatively general class of models, that includes the most prominent models in neuroscience (reviewed in appendix A), the propagation of chaos property took place, and showed convergence of the mean-field equations towards mean-field equations of a new type, analogous of the classical McKean-Vlasov equations but including delayed interactions, spatial integration term and a singular spatio-temporal stochastic process: the spatially chaotic Brownian motion.

The question of the scale at which relevant phenomena occur is essential to the modeler. Descriptions coarser than our neural field limit, for instance those involving finite numbers of populations, correspond to cases where our measure  $\lambda$  is a sum of Dirac masses. This case can be seen as a particular case of the present analysis, and hence the propagation of chaos occurs and network equations converge towards mean-field equations that correspond to a finite system of delayed McKean-Vlasov equations. In contrast, scales finer than the neural field limit (taking for instance into account possible individual heterogeneities between neurons) are not covered by the analysis and seem relatively hard to understand. It is likely that the dynamics of such networks will be considerably distinct from that of networks in the neural field regime. The neural-field regime seems particularly well suited to describe the activity of large neuronal assemblies, since it was observed that population sizes are orders of magnitude larger than the total number of populations [16]. Moreover, it seems to be at the scale of biological recordings and phenomena such as the emergence of patterns of activity in the cortex. We illustrated how such an analysis could be rigorously developed in simple example in section 4. More relevant states may be analyzed with this model, since the usual heuristic equations that were successfully used in a number of situations [13] are compatible, in the zero noise limit, with our equations, and the rigorously derived model will shed new light on the role of noise in such neuronal systems, but also on the individual behaviors of neurons. For instance, the propagation of chaos property ensures that finite sets of neurons are independent in the neural field limit. This result contradicts the classical view considering that since neurons of the same population are highly connected and receive similar input, their activity shall be correlated. However, with recent experimental findings using high quality recordings [11, 26] showed that levels of correlations between two neurons (of the same population or not) were extremely small, way below what was usually considered. The propagation of chaos hence offers a universal explanation to this phenomenon.

A number of open questions remains widely open in the theoretical understanding of the behavior of neural fields and large-scale neural networks. For instance, a particularly interesting phenomenon is the plasticity of neuronal connections. Considered constant and homogeneous in the present manuscript, it happens that the synaptic coefficients describing pairwise interactions between neurons evolve, very slowly, as a function of the network activity and in particular as a function of the correlations between the activity of pairs of neurons. This kind of phenomena was never considered in the mathematical literature, and seems relatively rich. In particular, this mechanism can break the propagation of chaos property and yield weakly correlated states. This is a problem we are currently investigating.

## Appendix A: Neuron models

For the sake of completeness we quickly review in this appendix different classical neuron models motivating the present study. This appendix takes a mathematical viewpoint, is obviously very selective and lacunar. The interested reader will find more details in classical neuroscience textbooks, e.g. [21, 12]. Basically, neurons are electrically excitable cells whose activity, measured through the voltage of the cell (difference of electrical potential between the intracellular and extracellular domains), is governed by ionic transfers through specific proteins (ion voltage-gated ion channels) located on the cellular membrane. We present here detailed neuron models (appendix A.1) that approximate the biophysics of ion channels, and firing-rate models (appendix A.2) that reproduce qualitatively the dynamics of the firing rate of neurons and that are used in the application section 4.

### A.1. Hodgkin-Huxley and Fitzhugh-Nagumo models

Probably the most biologically relevant, versatile and precise neuron model is the Hodgkin-Huxley (HH) model [18]. This model describes the membrane potential  $v$  of a neurons as a function of the dynamics of several ionic currents that enter or exit the cells through voltage-gated channels. The mathematical description we choose here involves Langevin approximation of the random proportion open of ion channel (see e.g. [17] and references therein). The proportion of open channels satisfies in that model a stochastic differential equation:

$$dx_t = (A_x(v)(1-x) - B_x(v)x) dt + \sqrt{A_x(v)(1-x) + B_x(v)x} \chi(x) dW_t^x$$

where  $W_t^x$  are independent standard Brownian motions,  $A_x(v)$  and  $B_x(v)$  are smooth bounded functions accounting respectively for the opening and closing probability intensity of a given channel and  $\chi(x)$  is a function vanishing outside  $[0, 1]$  to ensure that the variables  $x$  remain in  $[0, 1]$  (since these variables describe proportions). Generally, three ionic currents (and channels) are considered: potassium ( $m$ ), calcium activation ( $n$ ) and inactivation ( $h$ ), and Ohmic leak current,  $I_L$  (carried by  $Cl^-$  ions). Considering that the neuron receives an external current composed of a deterministic part  $I(t)$  and a white noise with standard deviation  $\sigma_{ext}$ , the voltage is governed by the equation:

$$\begin{cases} Cdv_t &= \left( I(t) - \bar{g}_K n^4 (v - E_K) - \bar{g}_{Na} m^3 h (v - E_{Na}) - \bar{g}_L (v - E_L) \right) dt + \sigma_{ext} dW_t \\ dx_t &= \left( A_x(v)(1-x) - B_x(v)x \right) dt + \sigma_x(v, n) dW_t^x \quad x \in \{n, m, h\} \end{cases} \quad (12)$$

This model satisfies assumptions (H1) and (H4) used in the general theory, since though polynomial nonlinearities arise in the dynamics, the boundedness of the variables ( $n, m, h$ ) ensure Lipschitz continuity and linear growth. The assumption (H5) is not satisfied since the noise depends on the state of the neuron. This refinement does not make the proofs substantially more intricate as discussed in appendix B.

The HH model is often too complex for practical purposes and several reductions were proposed. A particularly interesting one is the Fitzhugh-Nagumo (FN) bidimensional model [15] capturing from the biological viewpoint the most prominent behaviors of the Hodgkin-Huxley model. From the mathematical viewpoint, it is important to specify this model since that

model does not satisfy assumptions (H1) and (H4), and motivates the additional mathematical developments of appendix B. This model describes the evolution of the membrane potential variable  $v$  and a slower recovery variable  $w$ , through the equations:

$$\begin{cases} dv_t = (P(v_t) - w_t + I) dt + \sigma_v dW_t^v \\ dw_t = a(b v_t - w_t) dt + \sigma_w dW_t^w \end{cases} \quad (13)$$

where  $P(v) = v(1 - v)(v - a)$ , generally chosen  $f(v) = v - v^3$ .

The state of the neuron  $X$  in our abstract model (1) in the HH model is given by  $(v, n, m, h)$  and for the FN model by  $(v, w)$  and their intrinsic dynamics is enclosed in the functions  $f$  and  $g$ .

The communication between neurons is maintained by two possible types of synapses: electrical or chemical. Electrical synapses, in charge of rapid and stereotype signal transmission, is done through direct contact of the intracellular domain of the two communicating cells through specialized protein structures called gap-junctions. The ions passively flow from one neuron to the other: the interaction is not delayed and the current produced by neuron  $j$  on neuron  $i$  is equal to  $J_{ij}(v_t^j - v_t^i)$  where  $J_{ij}$  is called the synaptic conductance (this defines our interaction function  $b$  in the abstract model (1)). When including the dependence on  $v_t^i$  in the drift function, the interaction function  $\sum_j J_{ij} v_t^j$  clearly satisfies assumption (H2) and (H3), and (H5) as soon as the dependence of  $J_{ij}$  with respect to space is sufficiently regular. The chemical synapse is the most common type of interconnection. When a spike is fired from a pre-synaptic neuron  $j$ , it is transported through the axons to the synaptic button where it is transmitted to neuron  $i$  through a complex process of release of neurotransmitter (from  $j$ ) binding to specific receptors on neuron  $i$ . The transmission takes a time  $\tau_{ij}$  of the order of a few millisecond. Similarly to HH ion channels dynamics, the proportion of open neurotransmitter channels  $y^i$  has the dynamics (see [9]):

$$dy_t^j = \left( AS(v^j)(1 - y^j(t)) - D y^j(t) \right) dt + \sigma_Y(v^j, y^j) dW_t^{j,y}.$$

with  $S$  is a smooth sigmoidal function. In our abstract model, the variable  $y^i$  is added to the state  $X^i$  of neuron  $i$  and the functions  $f$  and  $g$  take into account that dynamics. The synaptic current induced at time  $t$  on neuron  $i$  by the arrival of a spike from neuron  $j$  (fired at time  $t - \tau_{ij}$ ) is equal to  $J_{ij} y^j(t - \tau_{ij})(v^i(t) - v_{\text{rev}})$  governing our interaction function  $b$  clearly satisfying assumptions (H2) and (H3), and (H5) as soon as the dependence of  $J_{ij}$  with respect to space is sufficiently regular.

The synaptic efficacies  $J_{ij}$  of electrical or chemical synapses are given by the connectivity of the cells. Such functions are generally considered continuous functions  $J(r_i, r_j)$  depending on the population of  $i$  and  $j$ .

Putting all these elements together and assuming that all the parameters of the equations only depend on the neural populations of the cells involved, we can write the equation of a network of FN neurons with chemical synapses, external and synaptic noise:

$$\begin{cases} dv_t^i &= \left( P(v_t^i) + I^i(t) + \sum_{j=1, j \neq i}^N (J_{ij} y^j(t - \tau_{ij})(v_t^i - v_{\text{rev}})) \right) dt \\ dw_t^i &= a_\alpha (b_\alpha v_t^i - w_t^i) dt \\ dy_t^i &= (A_\alpha S_\alpha(v_t^i)(1 - y_t^i) - D_\alpha y_t^i) dt + \sigma_Y(v, y) dB_t^{i,Y} \end{cases} \quad (14)$$

A similarly (but more complex) expression is obtained for the HH model using equations (12) and with distributed delays.



## A.2. Stochastic firing-rates models

A phenomenological neuron model consists in considering that neurons interact through their mean firing-rate. The firing-rate model considers that the membrane potential has a linear dynamics, and its mean-firing rate is a smooth sigmoidal transform of the membrane potential  $S(r_\alpha, \cdot)$  depending on the neural population  $\alpha$ . In other words, an incoming firing rate provokes postsynaptic potentials that linearly sum. The neurons receive additional inputs that are the sum of a deterministic current  $I(r_\alpha, t)$  and noise  $\sigma(r_\alpha)dW_t^i$ . The network equations hence read:

$$dV^i(t) = \left( -\frac{1}{\theta(r_\alpha)} V^i(t) + I(r_\alpha, t) + \sum_{\gamma=1}^P J_{\alpha\gamma} \frac{1}{N_\gamma} \sum_{j, p(j)=\gamma} S(r_\gamma, V^j(t - \tau_{\alpha\gamma})) \right) dt + \sigma(r_\alpha) dW_t^i$$

It is easy to check that assumptions (H1)-(H5) are satisfied for the firing-rate model.

## Appendix B: Generalized models

In the main section we chose to concentrate on the cornerstone mathematical problems arising in the modeling of neural fields, and chose to deal with relatively general models, yet simplified. Indeed, as discussed in section A, we saw that two technicalities were not taken into account in our general analysis. These were (i) non-globally Lipschitz drift that do not satisfy the linear growth condition (for Fitzhugh-Nagumo models) and (ii) state-dependent diffusion coefficients. Relatively classical methods allow to extend our proofs to these two refinements. In this section we will explain how one can extend to present analysis to models including this kind of dynamics. We consider the network equations:

$$dX_t^{i,N} = \left( f(r_\alpha, t, X_t^{i,N}) + \frac{1}{P(N)} \sum_{\gamma=1}^{P(N)} \sum_{p(j)=\gamma} \frac{1}{N_\gamma} b(r_\alpha, r_\gamma, X_t^{i,N}, X_{t-\tau(r_\alpha, r_\gamma)}^{j,N}) \right) dt + g(r_\alpha, t, X_t^{i,N}) dW_t^i \quad (15)$$

hence make the following generalized assumptions:

- (H1')  $f(r, t, \cdot)$  and  $g(r, \cdot)$  are locally Lipschitz-continuous
- (H4')  $x^T f(r, t, x) + \frac{1}{2} |g(r, t, x)|^2 \leq K (1 + |x|^2)$  uniformly in  $(r, t)$

The main difficulty is the non-global Lipschitz continuity of the drift and diffusion functions. However, under assumption (H4'), we can show that any possible solution is in  $\mathcal{M}^2([-\tau, T], \mathbb{L}_\lambda^2(\Gamma))$ , thanks to the following lemma:

**Lemma 5.** *Let  $(\zeta_t^0(r)) \in \mathcal{M}^2([-\tau, 0], \mathbb{L}_\lambda^2(\Gamma))$  be an initial condition for the mean-field equation (3). Any possible solution  $(\bar{X}_t)_{t \in [-\tau, T]}$  of the equation (3) with initial condition  $\zeta^0$  and measurable law with respect to  $\mathcal{B}(\Gamma)$  is square integrable, in the sense:*

$$\sup_{t \in [-\tau, T]} \mathbb{E} [|\mathcal{E}_t X_t(r)|^2] \leq C(T) \quad (16)$$

where  $C(T)$  is a quantity depending on the horizon  $T$  and the parameters of the system.

*Proof.* The proof is based on the application of Itô's formula for the squared modulus of  $X_t$ , standard inequalities and Gronwall's lemma. In details, let  $X$  be a solution of the mean-field equations, and

$$\tau_n = \inf \{t > 0 ; \|X_t\|_{\mathbb{L}_\lambda^2(\Gamma)} > n\}.$$

Due to the non-standard nature of the equation, let us underline the fact that Itô's formula is valid, i.e. that for any  $r \in \Gamma$ ,  $t \mapsto X_t(r)$  is a semimartingale. By definition, it is clear that for any  $r \in \Gamma$  both  $X_{t+s}(r)$  and  $Z_{t+s}(r)$  are  $\mathcal{F}_t$  measurable for all  $s \in [-\tau, 0]$  since these are driven by a standard Brownian motion  $W_t(r)$ , implying that  $X_t(r)$  is the sum of a continuous adapted process of finite variation:

$$\int_0^t \left( f(r, s, X_s(r)) + \mathcal{E}_{r'}[\mathbb{E}_{\bar{Z}}[b(r, r', \bar{X}_s(r), \bar{Z}_{s-\tau(r, r')}(r'))]] \right) ds,$$

a continuous  $(\mathcal{F}_t, \mathbb{P})$ -local martingale which is a stochastic integral of a progressively measurable processes with respect to a Brownian motion,  $\int_0^t g(r, s, \bar{X}_s^\alpha) dW_s(r)$ .

We can therefore apply Itô's formula ( see e.g. [8, Chap. 4.5] in the context of delayed equations) to  $\mathbb{E} [|\bar{X}_{t \wedge \tau_n}|^2]$ , and obtain:

$$\begin{aligned} \mathcal{E}_r [\mathbb{E} [|\bar{X}_{s \wedge \tau_n}(r)|^2]] &= \mathbb{E} \left[ \mathcal{E}_r \left( |\zeta_0^0(r)|^2 + 2 \sup_{s \in [0, t]} \int_0^{s \wedge \tau_n} du \left\{ \bar{X}_u^T(r) f(r, u, \bar{X}_u(r)) + \frac{1}{2} |g(r, u, \bar{X}_u(r))|^2 \right. \right. \right. \\ &\quad \left. \left. \left. + \mathcal{E}_{r'} \left( \bar{X}_u^T(r) \mathbb{E}_{\bar{Z}}[b(r, r', \bar{X}_u(r), Z_{u-\tau(r, r')}(r'))] \right) \right\} \right) \right] \\ &\leq \mathbb{E} [\mathcal{E}_r [|\zeta_0^0(r)|^2]] + 2 \left( K + \sqrt{\tilde{K}} \right) \int_0^{t \wedge \tau_n} du \mathbb{E} [\mathcal{E}_r (1 + |\bar{X}_u(r)|^2)], \end{aligned}$$

yielding, using Gronwall's lemma:

$$\sup_{t \in [0, T]} \mathbb{E} [\mathcal{E} [|\bar{X}_{t \wedge \tau_n}|^2]] \leq \mathbb{E} [\mathcal{E}_r [|\zeta_0^0(r)|^2]] e^{K'T} =: C'(T)$$

with  $K' = 2(K + \sqrt{\tilde{K}})$ . This estimate is valid for any  $n$ , thus for  $n$  sufficiently large, the probability of  $|X_t|$  to exceed  $n$  prior to time  $T$  vanishes. Letting  $n$  go to infinity, we obtain:

$$\sup_{t \in [-\tau, T]} \mathbb{E} [\|\bar{X}_t\|_{\mathbb{L}_\lambda^2(\Gamma)}^2] \leq \max \left( \sup_{s \in [-\tau, 0]} \mathbb{E} [\|\zeta_s^0\|_{\mathbb{L}_\lambda^2(\Gamma)}^2], C'(T) \right) =: C(T).$$

□

The same result can be shown when considering the network equations. These results being proved, it is then possible to prove analogous versions of theorems 2 and 3. The state dependent diffusion function is easily controlled using Burkholder-Davis-Gundy theorem, and the proofs of theorems 2 and 3 are hence valid when considering the truncated drift and diffusion functions:

$$f_U(r, t, x) = \begin{cases} f(r, t, x) & |x| \leq U \\ f(r, t, Ux/|x|) & |x| > U \end{cases}$$

and

$$g_U(r, t, x) = \begin{cases} g(r, t, x) & |x| \leq U \\ g(r, t, Ux/|x|) & |x| > U \end{cases}$$

which are both globally Lipschitz-continuous, and from these results and lemma 5, it is folklore to extend these results to the original problem.

Let us for instance focus on the existence and uniqueness of solutions for  $f$  and  $g$  functions that are not globally Lipschitz. Denoting  $\bar{X}_U$  the unique solution to the truncated problem, and defining the stopping time  $\tau_U = \inf\{t \in [0, T], \|\bar{X}_U(t)\|_{\mathbb{L}^2_\lambda(\Gamma)} \geq U\}$ , it is easy to show that

$$\bar{X}_U(t) = \bar{X}_{U'}(t) \quad \text{if} \quad 0 \leq t \leq \tau_U, \quad U' \geq U,$$

implying that the sequence of stopping times  $\tau_U$  is increasing. Using proposition 5 which implies that the solution to (3) is almost surely bounded, ensures existence and uniqueness of solutions of the mean-field equations (3) by letting  $U$  go to infinity. A similar argument applies for the propagation of chaos property.

## Acknowledgments

The author warmly acknowledges Clément Mouhot for very interesting discussions on the links between this research and kinetic theory. A special acknowledgement goes to Philippe Robert for great discussions and advice on the content and writing of the paper. He also acknowledges Gerard Ben Arous and Marc Yor for interesting discussions and suggestions and partial funding from ERC grant #227747.

## References

- [1] S. Amari. Characteristics of random nets of analog neuron-like elements. *Syst. Man Cybernet. SMC-2*, 1972.
- [2] Javier Baladron, Diego Fasoli, Olivier Faugeras, and Jonathan Touboul. Mean field description of and propagation of chaos in recurrent multipopulation networks of hodgkin-huxley and fitzhugh-nagumo neurons. arXiv:1110.4294, 2011.
- [3] W.H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6):2112–2127, 1997.
- [4] P.C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45:033001, 2012.
- [5] Gyorgy Buzsaki. *Rhythms of the brain*. Oxford University Press, USA, 2004.
- [6] S. Coombes and M. R. Owen. Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Phys. Rev. Lett.*, 94(14), 2005.
- [7] Stephen Coombes and Carlo Laing. Delays in activity based neural networks. *Submitted to the Royal Society*, 2011.
- [8] G Da Prato and J Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge Univ Pr, 1992.
- [9] A Destexhe, ZF Mainen, and TJ Sejnowski. An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Computation*, 6(1):14–18
- [10] RL Dobrushin. Prescribing a system of random variables by conditional distributions. *Theory of Probability and its Applications*, 15, 1970.

- [11] A.S. Ecker, P. Berens, G.A. Keliris, M. Bethge, N.K. Logothetis, and A.S. Tolias. Decorrelated neuronal firing in cortical microcircuits. *science*, 327(5965):584, January 2010.
- [12] G. Bard Ermentrout and David Terman. *Foundations Of Mathematical Neuroscience*. Interdisciplinary Applied Mathematics. Springer, 2010.
- [13] GB Ermentrout and J.D. Cowan. Temporal oscillations in neuronal nets. *Journal of mathematical biology*, 7(3):265–280, 1979.
- [14] G.B. Ermentrout and J.D. Cowan. Large scale spatially organized activity in neural nets. *SIAM Journal on Applied Mathematics*, pages 1–21, 1980.
- [15] R. FitzHugh. *Mathematical models of excitation and propagation in nerve*, chapter 1. McGraw-Hill Companies, 1969.
- [16] Y. Fregnac, M. Blatow, JP Changeux, J. De Felipe, A. Lansner, W. Maass, DA Mc Cormick, CM Michel, H. Monyer, E. Szathmáry, and R. Yuste. Ups and downs in cortical computation. *Microcircuits: the interface between neurons and global brain function*, *The MIT Press*, 393–433, 2006.
- [17] J.H. Goldwyn, N.S. Imennov, M. Famulare, and E. Shea-Brown. Stochastic differential equation models for ion channel noise in hodgkin-huxley neurons. *Physical Review E*, 83(4):041908, 2011.
- [18] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.
- [19] D. H. Hubel, T. N. Wiesel, and M. P. Stryker. Anatomical demonstration of orientation columns in macaque monkey. *J. Comp. Neur.*, 177:361–380, 1978.
- [20] E.M. Izhikevich. Polychronization: Computation with spikes. *Neural Computation*, 18(2):245–282
- [21] E.R. Kandel, J.H. Schwartz, and T.M. Jessel. *Principles of Neural Science*. McGraw-Hill, 4th edition, 2000.
- [22] C.L. Laing, W.C. Troy, B. Gutkin, and G.B. Ermentrout. Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.*, 63(1):62–97, 2002.
- [23] X. Mao. A note on the lasalle-type theorems for stochastic differential delay equations. *Journal of mathematical analysis and applications*, 268(1):125–142
- [24] HP McKean Jr. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6), 1966.
- [25] V.B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120:701–722, 1997.
- [26] A. Renart, J. De la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K.D. Harris. The asynchronous state in cortical circuits. *science*, 327(5965):587, January 2010.
- [27] D Revuz and M Yor. *Continuous martingales and Brownian motion*. Springer Verlag, 1999.
- [28] P. Series, S. Georges, J. Lorenceau, and Y. Frégnac. Orientation dependent modulation of apparent speed: a model based on the dynamics of feed-forward and horizontal connectivity in v1 cortex. *Vision research*, 42(25):2781–2797, 2002.
- [29] AS Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *Journal of Functional Analysis*, 56(3):311–336, 1984.
- [30] AS Sznitman. Topics in propagation of chaos. *Ecole d’Eté de Probabilités de Saint-Flour XIX*, pages 165–251, 1989.
- [31] H Tanaka. Some probabilistic problems in the spatially homogeneous boltzmann equation.

*Theory and Application of Random Fields*, pages 258–267, 1983.

- [32] S. Thorpe, A. Delorme, and R. VanRullen. Spike based strategies for rapid processing. *Neural Networks*, 14:715–726, 2001.
- [33] Jonathan Touboul. On the dynamics of mean-field equations for stochastic neural fields with delays.
- [34] C. Villani. A review of mathematical topics in collisional kinetic theory. *Handbook of mathematical fluid dynamics*, 1:71–74 1874–5792, 2002.
- [35] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24, 1972.
- [36] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics*, 13(2):55–80, September 1973.